



coordinator@medconfidential.org

20th October 2017

Dear civil society members of the Partnership on AI,

medConfidential welcomes and supports civil society involvement in the Partnership; we are confident you will all make invaluable contributions. AI will hopefully continue to bring benefits commensurate with its promise, but in doing so, as the Partnership recognises, needs to tread carefully to raise public understanding and avoid controversy and subsequent public distrust.

We write to you now about a project which has now been thoroughly investigated, involving 1.6 million identifiable medical histories and one of the Partnership's founder members, Google DeepMind.¹ The concerns about public statements and process that it raises are relevant to the Partnership as a whole, and may coincide with some of your interests. We would of course be very happy to discuss details with you individually, at your convenience.

From this letter, we have omitted much detail of the rules around use of medical records in the UK - we are happy to discuss them with you if you wish, but your focus is more likely to be on nature and culture of the incident response, rather than the details.

While this incident – which has taken a year to get to this stage – relates to one hospital in North London, it could equally be about self-driving cars in San Francisco,² or whatever the next scandal might be. Each time history repeats itself, the price to be paid goes up.

We are sure you'll agree that the underlying principle is not AI development *or* privacy, but AI development *and* privacy.³ And that requires honesty, integrity, and truthfulness. We ask you to consider a broader question: **are these the actions you expect from corporate Partners?**

In any future incident, how do you want Partners to communicate their actions with the public and civil society? Is it, for example, expected that Partners will mislead journalists and the public? Is it acceptable to loudly proclaim “transparency” while hiding the very page of the “public” document that evidences demonstrable falsehood, hoping no one else has a copy?⁴

¹ According to many public statements, it did not involve AI or machine learning; except that it did:

<http://www.hra.nhs.uk/news/research-summaries/using-machine-learning-to-improve-prediction-of-aki-deterioration/>

² <https://www.theverge.com/2017/2/27/14698902/uber-self-driving-san-francisco-dmv-email-levandowski>

³ With apologies to the UK Information Commissioner: <https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2016/09/transparency-trust-and-progressive-data-protection/>

⁴ Where is the page numbered 12? <https://deepmind.com/documents/22/REDACTED%20-%20FULLY%20EXECUTED%20DeepMind%20RFL%20Services%20Agreement.pdf>

Background: DeepMind and the Royal Free Hospital

The project in question involved DeepMind copying of 1.6 million patients' hospital records from the Royal Free London NHS Foundation Trust from late 2015 onwards, and subsequent public debate in April 2016 that began after Freedom of Information Act requests showed the public statements were fundamentally misleading.⁵ In the face of public scrutiny, there was continued denial and obfuscation by Google DeepMind.⁶

We would not expect the Partnership to determine whether the project was lawful or not; the UK's National Data Guardian, and separately the Information Commissioner (data protection authority), has already done that - it wasn't.⁷

DeepMind had its project approved by the hospital because the 'Privacy Impact Assessment' made no mention of the vast swathes of patient data being copied.⁸ An academic article in an obscure journal⁹ received far more detailed response and scrutiny from DeepMind than it allows when it makes selective information public. That the company attacked academic inquiry simply because it didn't like the outcome is a chilling precedent for the Partnership as it encourages academic research into areas where it may not necessarily immediately like the answers.¹⁰

When information was published, such as the letter containing the NDG's finding was published by a TV news channel, DeepMind's PR response was aggressive and misleading - they initially claimed not to have seen the letter that was copied to them.¹¹ In a later story, they denied to the press¹² claims that DeepMind had itself previously made to regulators.¹³

A press squabble between a flailing PR flack and a privacy NGO armed with citations may be little more than an entertaining sideshow to your organisations, but it should be of concern that this event occurred at all. The Partnership on AI lauds transparency, requires public confidence to deliver on members' goals, and above all wants to avoid the public believing AI is as creepy as some other silicon valley models.

medConfidential has few concerns about the other bona fide research projects which we are aware that DeepMind is pursuing. AI-assisted diagnosis will provide benefits to patients,

⁵ Page 4: <https://medconfidential.org/wp-content/uploads/2017/09/2017-06-10-Deepmind-NDG.pdf>

⁶ There was a squabble about the name of Google DeepMind, as this controversial project press release did not mention Google:

<https://www.royalfree.nhs.uk/news-media/news/new-app-helping-to-improve-patient-care/> yet this positive write-up of a different project clearly does: <https://www.theguardian.com/global/2017/mar/14/googles-deepmind-makes-ai-program-that-can-learn-like-a-human>

⁷ We enclose details of that determination.

⁸ There is no indication that DeepMind were aware of the omission from the PIA – DeepMind indicated that they hadn't read it and there's no reason, in this catalogue of errors, that they would have thought to do so.

⁹ <http://www.cam.ac.uk/research/news/deepmind-royal-free-deal-is-cautionary-tale-for-healthcare-in-the-algorithmic-age>

¹⁰ <https://medconfidential.org/2017/medconfidential-comment-on-google-deepmind-briefing-on-an-academic-paper/>

¹¹ Due to formatting, the Cc list was on the third page; Deepmind claimed the letter had only 2 pages.

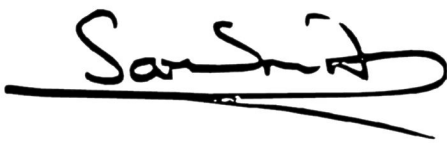
¹² "In addition, before any research could be done, DeepMind and the Royal Free would also need a research collaboration agreement," - Deepmind to the press

¹³ "We also have the Trust's full approval to conduct research on an anonymised copy of these data subject to ethical approval for which we hereby apply" - DeepMind to the Health Research Authority

wherever in the world they may be.¹⁴ Indeed, as a separate strand of work, we have published our thoughts on creating resilient public trust in technical environments which will be required as AI developers.¹⁵ We believe, for instance, that the work of the “Verifiable Data Audit” project at DeepMind is vital for AI safety at scale.

The work of the Partnership is also vital, but – if public trust is to be maintained – there must be consequences for “innovators” who mislead to the public, regulators, lawmakers, NGOs, or other interested parties. After all, it may be they are also lying to themselves.¹⁶ There will be inevitably be mistakes, and missteps; but these must be learned from transparently. The burning question in this letter – and in all future instances – is whether there will continue to be coverups without consequence.

Yours sincerely,



Sam Smith, medConfidential



Phil Booth, medConfidential

P.S. In our recent evidence to the UK’s House of Lords Select Committee on Artificial Intelligence,¹⁷ we said “*At the time of writing, Google DeepMind refuses to answer the question, “Did you feed the data to your AI?”*”. The latest published and citable answers¹⁸ in that area are “no AI has been applied to that dataset” (when it is not clear what “*that dataset*” refers to) and “No research project is underway” (no one said it was). We are still waiting for a very simple (non-caveated) answer, to a clear question (which is expressed here in a longer form, given DeepMind’s last non-responses): *Was any data from, or derived from, the Royal Free fed to an AI at any point?* In mid-2016, the public answer was a categorical “no” along with strong statement of the only lawful basis being direct care; 18 months later, we know that there was no lawful basis, and DeepMind’s answer to us and journalists has become “not in 2015” and “no research project is underway”. While we do understand DeepMind’s reason for such evasion,¹⁹ the question for the Partnership is whether this is how you want the partners that claim the “highest ethical standards” to act.²⁰

¹⁴ <https://medconfidential.org/2017/everyones-experience-in-ai-decision-making/>

¹⁵ Resilient public trust: <https://medconfidential.org/wp-content/uploads/2017/05/1-resilient-public-trust.pdf> & Modifying audit history <https://medconfidential.org/wp-content/uploads/2017/05/2-modifying-audit.pdf>

¹⁶ Para 57-58, medConfidential evidence to the House of Lords Select Committee on Artificial intelligence: <http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/artificial-intelligence-committee/artificial-intelligence/written/69500.html>

¹⁷

<http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/artificial-intelligence-committee/artificial-intelligence/written/69500.html>

¹⁸ <https://techcrunch.com/2017/08/31/documents-detail-deepminds-plan-to-apply-ai-to-nhs-data-in-2015/>

¹⁹ They have not yet passed a Privacy Impact Assessment and subsequent Audit, and recently claimed to be unaware of page 2 paragraph 2:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/633297/5_Ref_652-NDG_Letter_to_MHRA_FINAL_October_2016.pdf

²⁰ <https://deepmind.com/blog/why-we-launched-deepmind-ethics-society/>